Estimation and Evaluation of Conditional Asset Pricing Models
Paper by Stefan Nagel and Kenneth J. Singleton

Stephen Goldberger

University of North Carolina at Chapel Hill

February 8, 2010
Outline

Abstract
CRRA Model Example
Optimal GMM
  Moments
  GMM estimator
  Advantages
Optimal Portfolio Weights
  The Wald Test
  Maximal Power
  Portfolio Weights
  Results
Conclusions
Contributions of Paper

1. Construction of Optimal GMM estimator for SDF which is conditionally affine function of risk factors $S_{t} = E_{t}[m_{t+1} S_{t+1}]$.

2. Choosing the Optimal choice of Portfolios for testing estimated parameters against some null model of interest involving the SDF $H_{0}: m_{t+1} = a + b R_{m}$. 
Abstract

Contributions of Paper

1. Construction of Optimal GMM estimator for SDF which is conditionally affine function of risk factors

\[ S_t = \mathbb{E}_t [m_{t+1} S_{t+1}] \]
\[ m_{t+1} = a + b R_m + c \Delta y_t \]
Abstract

Contributions of Paper

1. Construction of Optimal GMM estimator for SDF which is conditionally affine function of risk factors

\[ S_t = \mathbb{E}_t [m_{t+1}S_{t+1}] \]
\[ m_{t+1} = a + bR_m + c\Delta y_t \]

2. Choosing the Optimal choice of Portfolios for testing estimated parameters against some null model of interest involving the SDF

\[ H_0 : m_{t+1} = a + bR_m \]
Example: Intertemporal Choice under uncertainty

Consumer’s Problem

\[
\max \sum_{h=0}^{\infty} \beta^h \mathbb{E}_t[u(c_{t+h})]
\]

s.t. \( c_t + S_t a_{t+1} = S_t a_t + d_t \)

First Order Conditions \([a_{t+1}]\)

\[
0 = -\mathbb{E}_t[S_t u'(c_t)] + \mathbb{E}_t[\beta S_{t+1} u'(c_{t+1})]
\]

\[
0 = -S_t u'(c_t) + \mathbb{E}_t[\beta S_{t+1} u'(c_{t+1})]
\]

\[
S_t = \mathbb{E}_t \left[ \beta \frac{u'(c_{t+1})}{u'(c_t)} S_{t+1} \right]
\]
CRRA Utility

\[ u(c_t) = \frac{c_t^{1-\sigma}}{1 - \sigma} \]

\[ u'(c_t) = c_t^{-\sigma} \]
CRRA Utility

\[ u(c_t) = \frac{c_t^{1-\sigma}}{1-\sigma} \]
\[ u'(c_t) = c_t^{-\sigma} \]

Asset Pricing Model

\[ S_t = \mathbb{E}_t \left[ \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\sigma} S_{t+1} \right] \]
Log Linearization

\[ \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\sigma} \approx \beta (1 - \sigma (\log c_{t+1} - \log c_t)) \]

\[ = \beta - \beta \sigma (\Delta c_t) \]
Log Linearization

\[
\beta \left( \frac{c_{t+1}}{c_t} \right)^{-\sigma} \approx \beta (1 - \sigma (\log c_{t+1} - \log c_t)) \\
= \beta - \beta \sigma (\triangle c_t)
\]

SDF

\[
S_t = \mathbb{E}_t \left[ \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\sigma} S_{t+1} \right] \\
\approx \mathbb{E}_t \left[ (\beta - \beta \sigma (\triangle c_t)) S_{t+1} \right] \\
= \mathbb{E}_t [m_{t+1} S_{t+1}]
\]
Conditionally Affine Stochastic Discount Factor

\[ m_{t+1}^G(\theta_0) = \phi^0(z_t; \beta_0, \gamma_0) + \phi^f(z_t; \beta_0, \gamma_0)f_{t+1} \]

\[ z_t \in \mathcal{I}_t \]

\[ f_{t+1} \text{ is a vector of risk factors} \]
Conditionally Affine Stochastic Discount Factor

\[ m_{t+1}^G(\theta_0) = \phi^0(z_t; \beta_0, \gamma_0) + \phi^{f'}(z_t; \beta_0, \gamma_0)f_{t+1} \]

\[ z_t \in \mathcal{I}_t \]

\[ f_{t+1} \text{ is a vector of risk factors} \]

1. In the example, \( f_{t+1} = \Delta c_t \),
Conditionally Affine Stochastic Discount Factor

\[ m_{t+1}(\theta_0) = \phi^0(z_t; \beta_0, \gamma_0) + \phi^f(z_t; \beta_0, \gamma_0)f_{t+1} \]

\[ z_t \in \mathcal{I}_t \]

\[ f_{t+1} \text{ is a vector of risk factors} \]

1. In the example, \( f_{t+1} = \Delta c_t \),

2. Affine structure comes up from classical CAPM (Betas of Portfolios) and multifactor tweaks as well
Conditionally Affine Stochastic Discount Factor

\[ m_{t+1}^G(\theta_0) = \phi^0(z_t; \beta_0, \gamma_0) + \phi^f(z_t; \beta_0, \gamma_0)f_{t+1} \]

\[ z_t \in I_t \]

\[ f_{t+1} \text{ is a vector of risk factors} \]

1. In the example, \( f_{t+1} = \triangle c_t \),
2. Affine structure comes up from classical CAPM (Betas of Portfolios) and multifactor tweaks as well
3. As constructed, let \( m_{t+1}^N(\theta_0) \) be a null model of interest, with \( \gamma_0 = 0 \)
Conditionally Affine Stochastic Discount Factor

\[ m_{t+1}^G(\theta_0) = \phi^0(z_t; \beta_0, \gamma_0) + \phi^{f'}(z_t; \beta_0, \gamma_0)f_{t+1} \]

\[ z_t \in I_t \]

\[ f_{t+1} \text{ is a vector of risk factors} \]

1. In the example, \( f_{t+1} = \Delta c_t \),

2. Affine structure comes up from classical CAPM (Betas of Portfolios) and multifactor tweaks as well

3. As constructed, let \( m_{t+1}^N(\theta_0) \) be a null model of interest, with \( \gamma_0 = 0 \)

4. How to best estimate \( \theta_0 \) for SDFs of this type?
Optimal GMM Estimator

Moment Conditions
Define the pricing errors: \( h_{t+1}(\theta_0) \) as follows:

\[
h_{t+1}(\theta_0) = m_{t+1}(\theta_0)S_{t+1} - S_t
\]

Assume our model states that:

\[
E_t[m_{t+1}(\theta_0)S_{t+1} - S_t] = 0
\]
\[
E_t[h_{t+1}(\theta_0)] = 0
\]

This says that Pricing Errors are a Martingale Difference Sequence. As such we know that \( h_{t+1}(\theta_0) \) is orthogonal to all \( z_t \) in the information set \( \mathcal{I}_t \) for any given asset.
Moment Conditions
Given $K$ unknown parameters for $\theta_0$ and $R$ assets:
We can use the fact that pricing errors are orthogonal to 
Information at time $t$, and write our Moment condition

$$\mathbb{E}[A_t h_{t+1}(\theta_0)] = 0$$

$A_t$ is a $K \times R$ matrix of entries in $\mathcal{I}_t$
$h_{t+1}(\theta_0)$ is a $R \times 1$ vector of asset pricing errors.

This gives us $K$ equations, for $K$ unknowns.
**GMM**

The GMM estimator $\hat{\theta}_t^A$ is

$$\arg\min_{\theta} \left( \frac{1}{T} \sum_{t=1}^{T} A_t h_{t+1}(\theta) \right)' W \left( \frac{1}{T} \sum_{t=1}^{T} A_t h_{t+1}(\theta) \right)$$

for any given instruments $A_t$
GMM
The GMM estimator $\hat{\theta}_t^A$ is

$$\arg\min_{\theta} \left( \frac{1}{T} \sum_{t=1}^{T} A_t h_{t+1}(\theta) \right)' W \left( \frac{1}{T} \sum_{t=1}^{T} A_t h_{t+1}(\theta) \right)$$

for any given instruments $A_t$

Asymptotic Covariance Matrix
Using the optimal weighting matrix
$W = \Sigma^A_0 = \mathbb{E}[A_t h_{t+1}(\theta_0) h_{t+1}(\theta_0)' A_t]^{-1}$

$$\Omega^A_0 = \mathbb{E} \left[ A_t \frac{\partial h_{t+1}(\theta)}{\partial \theta} \right]^{-1} \Sigma^A_0 \mathbb{E} \left[ \frac{\partial h_{t+1}(\theta)'}{\partial \theta} A_t \right]^{-1}$$
Valid Estimators

Any Instruments $A_t$ are valid as long as:

$$A_t \text{ is known at time } t$$

$$E_h A_t \partial h_{t+1} \partial i \text{ has full rank}$$

Question

Which instruments do we choose in order to minimize $\Omega A_0$ and thus have the most efficient estimator of $\Omega$?
Valid Estimators

Any Instruments $A_t$ are valid as long as:

1. $A_t$ is known at time $t$
Valid Estimators

Any Instruments $A_t$ are valid as long as:

1. $A_t$ is known at time $t$
2. $\mathbb{E} \left[ A_t \frac{\partial h_{t+1}(\theta)}{\partial \theta} \right]$ has full rank
Valid Estimators
Any Instruments $A_t$ are valid as long as:

1. $A_t$ is known at time $t$
2. $\mathbb{E} \left[ A_t \frac{\partial h_{t+1}(\theta)}{\partial \theta} \right]$ has full rank

Question
Which instruments do we choose in order to minimize $\Omega^A_0$ and thus have the most efficient estimator of $\theta_0$?
Efficient GMM for Martingale Difference Sequences

Optimal choice for $A_t$ is

$$A_t^* = \Psi_t^\theta \Sigma_t^{-1}$$

where $\Psi_t^\theta \equiv \mathbb{E}_t\left[\frac{\partial h_{t+1}(\theta)}{\partial \theta}\right]$.

Corresponding asymptotic covariance matrix:

$$\Omega_0^* = (\mathbb{E}[\Psi_t^\theta \Sigma_t^{-1} \Psi_t^\theta])^{-1}$$

Intuitively, this instrument captures the sensitivity of pricing errors to changes to parameters.

In general \( A_t^* = \psi_t^\theta \Sigma_t^{-1} \notin \mathcal{I}_t \)
Problem
In general $A^*_t = \Psi_t' \Sigma_t^{-1} \notin \mathcal{I}_t$

However
In the case where $m_{t+1}(\theta)$ is an affine function of conditional risk factors this instrument has a clear functional form:

$$
\mathbb{E}_t \left[ \frac{\partial h_{t+1}(\theta_0)}{\partial \theta} \right] = \frac{\partial \tilde{\phi}(z_t, \theta_0)'}{\partial \theta_0 j} \mathbb{E}_t[\tilde{f}_{t+1} r_{i,t+1}]
$$

The partial derivative is known at time $t$, $\mathbb{E}_t[\tilde{f}_{t+1} r_{i,t+1}]$ must be computed
Conditional Moments
How to Estimate $E_t[\tilde{f}_{t+1} r_{i,t+1}]$:
Conditional Moments
How to Estimate $E_t[\tilde{f}_{t+1} r_{i,t+1}]$:

1. Amass a vector of instruments $z_t$ at time $t$
Conditional Moments

How to Estimate $E_t[\tilde{f}_{t+1}r_{i,t+1}]$:

1. Amass a vector of instruments $z_t$ at time $t$
2. Local Linear Regression of $\tilde{f}_{t+1}r_{i,t+1}$ on $z_t$
Conditional Moments

How to Estimate $E_t[\tilde{f}_{t+1}r_{i,t+1}]$:

1. Amass a vector of instruments $z_t$ at time $t$
2. Local Linear Regression of $\tilde{f}_{t+1}r_{i,t+1}$ on $z_t$
3. Get $\hat{g}(z_t)$ as an estimate for conditional expectation
Conditional Moments
How to Estimate $E_t[\tilde{f}_{t+1}r_{i,t+1}]$:

1. Amass a vector of instruments $z_t$ at time $t$
2. Local Linear Regression of $\tilde{f}_{t+1}r_{i,t+1}$ on $z_t$
3. Get $\hat{g}(z_t)$ as an estimate for conditional expectation
4. Or an alternate strategy is to use a sieve method, where $z_t$ is cut down to ”optimal regressors” chosen by the Akaike Information Criterion (AIC)
Advantages

GMM estimator based on conditional expectations

1. Allow for State Dependent Factor-Betas and Market Prices of Risk
2. Use of the fact that $S_t = E_t [m_{t+1} S_{t+1}]$ implies if plug in the risk free asset: $E_t [m_{t+1}] = r_f t$. 
3. Conditional Expectations then more revealing about strengths and weaknesses of SDFs as descriptions of history

Stephen Goldberger
Advantages

GMM estimator based on conditional expectations

1. Allow for State Dependent Factor-Betas and Market Prices of Risk
Advantages

GMM estimator based on conditional expectations

1. Allow for State Dependent Factor-Betas and Market Prices of Risk

2. Use of the fact that \( S_t = \mathbb{E}_t[m_{t+1}S_{t+1}] \) implies if plug in the risk free asset: \( \mathbb{E}_t[m_{t+1}] = \frac{1}{r_t} \).
Advantages

GMM estimator based on conditional expectations

1. Allow for State Dependent Factor-Betas and Market Prices of Risk

2. Use of the fact that $S_t = \mathbb{E}_t[m_{t+1}S_{t+1}]$ implies if plug in the risk free asset: $\mathbb{E}_t[m_{t+1}] = \frac{1}{r_t}$.

3. Conditional Expectations then more revealing about strengths and weaknesses of SDFs as descriptions of history
The Wald Test

Testing

We hope to evaluate the goodness of fit for $m_{t+1}^G(\beta_0, \gamma_0)$ vs. a null $m_{t+1}^N(\beta_0, 0)$. 
The Wald Test

Testing
We hope to evaluate the goodness of fit for \( m_{t+1}^{G}(\beta_0, \gamma_0) \) vs. a null \( m_{t+1}^{N}(\beta_0, 0) \).

Under the Null
For any Instruments \( A_t \in I_t \)

\[
\varsigma_T^W (A) \equiv T \gamma_T' (\Omega^A_{\gamma \gamma})^{-1} \gamma_T
\]

\[
\varsigma_T^W (A) \to \chi^2(G)
\]

where \( G \) is the dimension of \( \gamma_0 \)
Maximum Power
We seek to maximize the probability of rejection of Null when it is false.
Maximum Power

We seek to maximize the probability of rejection of Null when it is false.

1. Under Null, compared to local alternatives, our choice of optimal instruments gives us maximal power among all other instruments.
Maximum Power
We seek to maximize the probability of rejection of Null when it is false.

1. Under Null, compared to local alternatives, our choice of optimal instruments gives us maximal power among all other instruments

2. Intuitively, this is because our instrument adjusts for sensitivity to price error to changes in choice of $\tilde{\theta}$
Portfolio Choice

Asymptotic Equivalence

Asymptotically: \( \sqrt{W_T}(A^*) = \)

\[
\left( \frac{1}{\sqrt{T}} \sum_{t=1}^{T} h_{t+1}(\theta_0)' \Sigma_{t}^{G-1} \mathcal{H}_t^{G} \right) \Omega_{\gamma\gamma}^* \left( \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \mathcal{H}_t^{G'} \Sigma_{t}^{G-1} h_{t+1}(\theta_0) \right)
\]

where

\[
\Psi_t^{\gamma} \equiv \mathbb{E}_t \left[ \frac{\partial h_{t+1}(\beta_0, \gamma_0)}{\partial \gamma} \right] , \quad \Psi_t^{\beta} \equiv \mathbb{E}_t \left[ \frac{\partial h_{t+1}(\beta_0, \gamma_0)}{\partial \beta} \right] \\
\mathcal{K}^{\beta\gamma} \equiv \mathbb{E} \left[ \Psi_t^{\beta'} \Sigma_t^{-1} \Psi_t^{\gamma} \right] , \quad \mathcal{H}_t \equiv \Psi_t^{\gamma} - \Psi_t^{\beta} (\mathcal{K}^{\beta\beta})^{-1} \mathcal{K}^{\beta\gamma}
\]
Equivalently

Wald test can be seen locally as a test of:

$$E[h_t^G \Sigma_t^{G^{-1}} h_t+1(\theta_0)] = 0$$

or whether managed portfolio returns $h_t^G \Sigma_t^{G^{-1}} r_{t+1}$ are priced correctly by $m_t^G$. 

Equivalently

Wald test can be seen locally as a test of:

\[ \mathbb{E}[H_t^G \Sigma_t^{G-1} h_{t+1}(\theta_0)] = 0 \]

or whether managed portfolio returns \( H_t^G \Sigma_t^{G-1} r_{t+1} \) are priced correctly by \( m_t^G \).

Factoring

Rewriting \( \Sigma_t^{-1} = D_t^{-1/2} D_t^{-1/2} \), the null \( \gamma_0 = 0 \) specifies that

\[ \mathbb{E}[D_t^{-1/2} \Psi_t^\beta \Sigma_t^{G-1} h_{t+1}(\theta_0)] = 0 \]

The Wald test uses the part of \( D_t^{-1/2} \Psi_t^\gamma \) that is orthogonal to this information.
Similarly
These portfolio weights can be derived using Lagrange Multiplier Test specification
Similarly
These portfolio weights can be derived using Lagrange Multiplier Test specification

Equivalently
Wald and LM test can be seen as whether or not $m_{t+1}^G$ accurately prices previously defined portfolio returns vs. $m_{t+1}^N$
\textbf{Implimentation}

Applying this GMM estimator and Test based on Conditional Moments, finds several models with Affine Stochastic Factors fail when taken to data. This is despite the fact that they seemingly provide accurate readings of the risk premium when considering GMM estimation based on unconditional moments.
Conclusions

With regard to Models with SDFs which are linear functions of factors
Conclusions

With regard to Models with SDFs which are linear functions of factors

1. There is an Optimal GMM estimator using Conditional Expectations
Conclusions

With regard to Models with SDFs which are linear functions of factors

1. There is an Optimal GMM estimator using Conditional Expectations

2. There is an Optimal choice for Portfolio Weights when testing specified SDF v. a Null SDF
Conclusions
With regard to Models with SDFs which are linear functions of factors

1. There is an Optimal GMM estimator using Conditional Expectations
2. There is an Optimal choice for Portfolio Weights when testing specified SDF v. a Null SDF
3. When you apply these techniques to Models in the Literature, you receive different results
Conclusions

With regard to Models with SDFs which are linear functions of factors

1. There is an Optimal GMM estimator using Conditional Expectations

2. There is an Optimal choice for Portfolio Weights when testing specified SDF v. a Null SDF

3. When you apply these techniques to Models in the Literature, you receive different results
Conclusions

With regard to Models with SDFs which are linear functions of factors

1. There is an Optimal GMM estimator using Conditional Expectations
2. There is an Optimal choice for Portfolio Weights when testing specified SDF v. a Null SDF
3. When you apply these techniques to Models in the Literature, you receive different results
   b) This is because small average pricing errors based on unconditional moment restrictions hide very large time variation in conditional pricing errors